

Sequence Listing

<110> Chou, Min-Yuan
Leu, Charng-Yih

<120> Novel Human alpha 1 Chain Collagen

<130> 32350-176844

<140> US 09/996,611

<141> 2001-11-30

<150> TAIWAN 89128027

<151> 2000-12-27

<160> 5

<170> Microsoft Word

<210> 1

<211> 954

<212> PRT

<213> Homo sapiens

<220>

<223> alpha 1 chain collagen

Pro in this sequence stands for hydroxyproline

<400> 1

Met	Ala	His	Tyr	Ile	Thr	Phe	Leu	Cys	Met	Val	Leu	Val	Leu	Leu
1				5					10					15

Leu	Gln	Asn	Ser	Val	Leu	Ala	Glu	Asp	Gly	Glu	Val	Arg	Ser	Ser
				20					25					30

Cys	Arg	Thr	Ala	Pro	Thr	Asp	Leu	Val	Phe	Ile	Leu	Asp	Gly	Ser
				35					40					45

Tyr	Ser	Val	Gly	Pro	Glu	Asn	Phe	Glu	Ile	Val	Lys	Lys	Trp	Leu
				50					55					60

Val	Asn	Ile	Thr	Lys	Asn	Phe	Asp	Ile	Gly	Pro	Lys	Phe	Ile	Gln
				65					70					75

Val	Gly	Val	Val	Gln	Tyr	Ser	Asp	Tyr	Pro	Val	Leu	Glu	Ile	Pro
				80					85					90

Leu	Gly	Ser	Tyr	Asp	Ser	Gly	Glu	His	Leu	Thr	Ala	Ala	Val	Glu
				95					100					105

Sequence Listing

Ser	Ile	Leu	Tyr	Leu	Gly	Gly	Asn	Thr	Lys	Thr	Gly	Lys	Ala	Ile	110	115	120
Gln	Phe	Ala	Leu	Asp	Tyr	Leu	Phe	Ala	Lys	Ser	Ser	Arg	Phe	Leu	125	130	135
Thr	Lys	Ile	Ala	Val	Val	Leu	Thr	Asp	Gly	Lys	Ser	Gln	Asp	Asp	140	145	150
Val	Lys	Asp	Ala	Ala	Gln	Ala	Ala	Arg	Asp	Ser	Lys	Ile	Thr	Leu	155	160	165
Phe	Ala	Ile	Gly	Val	Gly	Ser	Glu	Thr	Glu	Asp	Ala	Glu	Leu	Arg	170	175	180
Ala	Ile	Ala	Asn	Lys	Pro	Ser	Ser	Thr	Tyr	Val	Phe	Tyr	Val	Glu	185	190	195
Asp	Tyr	Ile	Ala	Ile	Ser	Lys	Ile	Arg	Glu	Val	Met	Lys	Gln	Lys	200	205	210
Leu	Cys	Glu	Glu	Ser	Val	Cys	Pro	Thr	Arg	Ile	Pro	Val	Ala	Ala	215	220	225
Arg	Asp	Glu	Arg	Gly	Phe	Asp	Ile	Leu	Leu	Gly	Leu	Asp	Val	Asn	230	235	240
Lys	Lys	Val	Lys	Lys	Arg	Ile	Gln	Leu	Ser	Pro	Lys	Lys	Ile	Lys	245	250	255
Gly	Tyr	Glu	Val	Thr	Ser	Lys	Val	Asp	Leu	Ser	Glu	Leu	Thr	Ser	260	265	270
Asn	Val	Phe	Pro	Glu	Gly	Leu	Pro	Pro	Ser	Tyr	Val	Phe	Val	Ser	275	280	285
Thr	Gln	Arg	Phe	Lys	Val	Lys	Lys	Ile	Trp	Asp	Leu	Trp	Arg	Ile	290	295	300
Leu	Thr	Ile	Asp	Gly	Arg	Pro	Gln	Ile	Ala	Val	Thr	Leu	Asn	Gly	305	310	315
Val	Asp	Lys	Ile	Leu	Leu	Phe	Thr	Thr	Thr	Ser	Val	Ile	Asn	Gly	320	325	330
Ser	Gln	Val	Val	Thr	Phe	Ala	Asn	Pro	Gln	Val	Lys	Thr	Leu	Phe	335	340	345

Sequence Listing

Asp	Glu	Gly	Trp	His	Gln	Ile	Arg	Leu	Leu	Val	Thr	Glu	Gln	Asp	350	355	360
Val	Thr	Leu	Tyr	Ile	Asp	Asp	Gln	Gln	Ile	Glu	Asn	Lys	Pro	Leu	365	370	375
His	Pro	Val	Leu	Gly	Ile	Leu	Ile	Asn	Gly	Gln	Thr	Gln	Ile	Gly	380	385	390
Lys	Tyr	Ser	Gly	Lys	Glu	Glu	Thr	Val	Gln	Phe	Asp	Val	Gln	Lys	395	400	405
Leu	Arg	Ile	Tyr	Cys	Asp	Pro	Glu	Gln	Asn	Asn	Arg	Glu	Thr	Ala	410	415	420
Cys	Glu	Ile	Pro	Gly	Phe	Cys	Leu	Asn	Gly	Pro	Ser	Asp	Val	Gly	425	430	435
Ser	Thr	Pro	Ala	Pro	Cys	Ile	Cys	Pro	Pro	Gly	Lys	Pro	Gly	Leu	440	445	450
Gln	Gly	Pro	Lys	Gly	Asp	Pro	Gly	Leu	Pro	Gly	Asn	Pro	Gly	Tyr	455	460	465
Pro	Gly	Gln	Pro	Gly	Gln	Asp	Gly	Lys	Pro	Gly	Tyr	Gln	Gly	Ile	470	475	480
Ala	Gly	Thr	Pro	Gly	Val	Pro	Gly	Ser	Pro	Gly	Ile	Gln	Gly	Ala	485	490	495
Arg	Gly	Leu	Pro	Gly	Tyr	Lys	Gly	Glu	Pro	Gly	Arg	Asp	Gly	Asp	500	505	510
Lys	Gly	Asp	Arg	Gly	Leu	Pro	Gly	Phe	Pro	Gly	Leu	His	Gly	Met	515	520	525
Pro	Gly	Ser	Lys	Gly	Glu	Met	Gly	Ala	Lys	Gly	Asp	Lys	Gly	Ser	530	535	540
Pro	Gly	Phe	Tyr	Gly	Lys	Lys	Gly	Ala	Lys	Gly	Glu	Lys	Gly	Asn	545	550	555
Ala	Gly	Phe	Pro	Gly	Leu	Pro	Gly	Pro	Ala	Gly	Glu	Pro	Gly	Arg	560	565	570
His	Gly	Lys	Asp	Gly	Leu	Met	Gly	Ser	Pro	Gly	Phe	Lys	Gly	Glu	575	580	585

Sequence Listing

Ala Gly Ser Pro Gly	Ala Pro Gly Gln Asp Gly Thr Arg Gly Glu	590	595	600
Pro Gly Ile Pro Gly Phe Pro Gly Asn Arg Gly Leu Met Gly Gln		605	610	615
Lys Gly Glu Ile Gly Pro Pro Gly Gln Gln Gly Lys Lys Gly Ala		620	625	630
Pro Gly Met Pro Gly Leu Met Gly Ser Asn Gly Ser Pro Gly Gln		635	640	645
Pro Gly Thr Pro Gly Ser Lys Gly Ser Lys Gly Glu Pro Gly Ile		650	655	660
Gln Gly Met Pro Gly Ala Ser Gly Leu Lys Gly Glu Pro Gly Ala		665	670	675
Thr Gly Ser Pro Gly Glu Pro Gly Tyr Met Gly Leu Pro Gly Ile		680	685	690
Gln Gly Lys Lys Gly Asp Lys Gly Asn Gln Gly Glu Lys Gly Ile		695	700	705
Gln Gly Gln Lys Gly Glu Asn Gly Arg Gln Gly Ile Pro Gly Gln		710	715	720
Gln Gly Ile Gln Gly His His Gly Ala Lys Gly Glu Arg Gly Glu		725	730	735
Lys Gly Glu Pro Gly Val Arg Gly Ala Ile Gly Ser Lys Gly Glu		740	745	750
Ser Gly Val Asp Gly Leu Met Gly Pro Ala Gly Pro Lys Gly Gln		755	760	765
Pro Gly Asp Pro Gly Pro Gln Gly Pro Pro Gly Leu Asp Gly Lys		770	775	780
Pro Gly Arg Glu Phe Ser Glu Gln Phe Ile Arg Gln Val Cys Thr		785	790	795
Asp Val Ile Arg Ala Gln Leu Pro Val Leu Leu Gln Ser Gly Arg		800	805	810
Ile Arg Asn Cys Asp His Cys Leu Ser Gln His Gly Ser Pro Gly		815	820	825

Sequence Listing

Ile Pro Gly Pro Pro Gly Pro Ile Gly Pro Glu Gly Pro Arg Gly
830 835 840

Leu Pro Gly Leu Pro Gly Arg Asp Gly Val Pro Gly Leu Val Gly
845 850 855

Val Pro Gly Arg Pro Gly Val Arg Gly Leu Lys Gly Leu Pro Gly
860 865 870

Arg Asn Gly Glu Lys Gly Ser Gln Gly Phe Gly Tyr Pro Gly Glu
875 880 885

Gln Gly Pro Pro Gly Pro Pro Gly Pro Glu Gly Pro Pro Gly Ile
890 895 900

Ser Lys Glu Gly Pro Pro Gly Asp Pro Gly Leu Pro Gly Lys Asp
905 910 915

Gly Asp His Gly Lys Pro Gly Ile Gln Gly Gln Pro Gly Pro Pro
920 925 930

Gly Ile Cys Asp Pro Ser Leu Cys Phe Ser Val Ile Ala Arg Arg
935 940 945

Asp Pro Phe Arg Lys Gly Pro Asn Tyr
950

<210> 2

<211> 171

<212> PRT

<213> Homo sapiens

<220>

<223> von Willebrand factor A domain

<400> 2

Asp Leu Val Phe Ile Leu Asp Gly Ser Tyr Ser Val Gly Pro Glu
1 5 10 15

Asn Phe Glu Ile Val Lys Lys Trp Leu Val Asn Ile Thr Lys Asn
20 25 30

Phe Asp Ile Gly Pro Lys Phe Ile Gln Val Gly Val Val Gln Tyr
35 40 45

Ser Asp Tyr Pro Val Leu Glu Ile Pro Leu Gly Ser Tyr Asp Ser
50 55 60

Sequence Listing

Gly Glu His Leu Thr Ala Ala Val Glu Ser Ile Leu Tyr Leu Gly
65 70 75

Gly Asn Thr Lys Thr Gly Lys Ala Ile Gln Phe Ala Leu Asp Tyr
80 85 90

Leu Phe Ala Lys Ser Ser Arg Phe Leu Thr Lys Ile Ala Val Val
95 100 105

Leu Thr Asp Gly Lys Ser Gln Asp Asp Val Lys Asp Ala Ala Gln
110 115 120

Ala Ala Arg Asp Ser Lys Ile Thr Leu Phe Ala Ile Gly Val Gly
125 130 135

Ser Glu Thr Glu Asp Ala Glu Leu Arg Ala Ile Ala Asn Lys Pro
140 145 150

Ser Ser Thr Tyr Val Phe Tyr Val Glu Asp Tyr Ile Ala Ile Ser
155 160 165

Lys Ile Arg Glu Val Met
170

<210> 3

<211> 183

<212> PRT

<213> Homo sapiens

<220>

<223> Thrombospondin N-terminal-like domain

<400> 3

Gly Phe Asp Ile Leu Leu Gly Leu Asp Val Asn Lys Lys Val Lys
1 5 10 15

Lys Arg Ile Gln Leu Ser Pro Lys Lys Ile Lys Gly Tyr Glu Val
20 25 30

Thr Ser Lys Val Asp Leu Ser Glu Leu Thr Ser Asn Val Phe Pro
35 40 45

Glu Gly Leu Pro Pro Ser Tyr Val Phe Val Ser Thr Gln Arg Phe
50 55 60

Lys Val Lys Lys Ile Trp Asp Leu Trp Arg Ile Leu Thr Ile Asp
65 70 75

Sequence Listing

Gly	Arg	Pro	Gln	Ile	Ala	Val	Thr	Leu	Asn	Gly	Val	Asp	Lys	Ile	80	85	90
Leu	Leu	Phe	Thr	Thr	Thr	Ser	Val	Ile	Asn	Gly	Ser	Gln	Val	Val	95	100	105
Thr	Phe	Ala	Asn	Pro	Gln	Val	Lys	Thr	Leu	Phe	Asp	Glu	Gly	Trp	110	115	120
His	Gln	Ile	Arg	Leu	Leu	Val	Thr	Glu	Gln	Asp	Val	Thr	Leu	Tyr	125	130	135
Ile	Asp	Asp	Gln	Gln	Ile	Glu	Asn	Lys	Pro	Leu	His	Pro	Val	Leu	140	145	150
Gly	Ile	Leu	Ile	Asn	Gly	Gln	Thr	Gln	Ile	Gly	Lys	Tyr	Ser	Gly	155	160	165
Lys	Glu	Glu	Thr	Val	Gln	Phe	Asp	Val	Gln	Lys	Leu	Arg	Ile	Tyr	170	175	180

Cys Asp Pro

<210> 4
 <211> 509
 <212> PRT
 <213> Homo sapiens

<220>
 <223> collagenous domain
 Pro in this sequence stands for hydroxyproline

<400>	4																
Gly	Lys	Pro	Gly	Leu	Gln	Gly	Pro	Lys	Gly	Asp	Pro	Gly	Leu	Pro			
1				5					10					15			
Gly	Asn	Pro	Gly	Tyr	Pro	Gly	Gln	Pro	Gly	Gln	Asp	Gly	Lys	Pro			
				20					25					30			
Gly	Tyr	Gln	Gly	Ile	Ala	Gly	Thr	Pro	Gly	Val	Pro	Gly	Ser	Pro			
				35					40					45			
Gly	Ile	Gln	Gly	Ala	Arg	Gly	Leu	Pro	Gly	Tyr	Lys	Gly	Glu	Pro			
				50					55					60			

Sequence Listing

Gly Arg Asp Gly Asp Lys Gly Asp Arg Gly Leu Pro Gly Phe Pro	65	70	75
Gly Leu His Gly Met Pro Gly Ser Lys Gly Glu Met Gly Ala Lys	80	85	90
Gly Asp Lys Gly Ser Pro Gly Phe Tyr Gly Lys Lys Gly Ala Lys	95	100	105
Gly Glu Lys Gly Asn Ala Gly Phe Pro Gly Leu Pro Gly Pro Ala	110	115	120
Gly Glu Pro Gly Arg His Gly Lys Asp Gly Leu Met Gly Ser Pro	125	130	135
Gly Phe Lys Gly Glu Ala Gly Ser Pro Gly Ala Pro Gly Gln Asp	140	145	150
Gly Thr Arg Gly Glu Pro Gly Ile Pro Gly Phe Pro Gly Asn Arg	155	160	165
Gly Leu Met Gly Gln Lys Gly Glu Ile Gly Pro Pro Gly Gln Gln	170	175	180
Gly Lys Lys Gly Ala Pro Gly Met Pro Gly Leu Met Gly Ser Asn	185	190	195
Gly Ser Pro Gly Gln Pro Gly Thr Pro Gly Ser Lys Gly Ser Lys	200	205	210
Gly Glu Pro Gly Ile Gln Gly Met Pro Gly Ala Ser Gly Leu Lys	215	220	225
Gly Glu Pro Gly Ala Thr Gly Ser Pro Gly Glu Pro Gly Tyr Met	230	235	240
Gly Leu Pro Gly Ile Gln Gly Lys Lys Gly Asp Lys Gly Asn Gln	245	250	255
Gly Glu Lys Gly Ile Gln Gly Gln Lys Gly Glu Asn Gly Arg Gln	260	265	270
Gly Ile Pro Gly Gln Gln Gly Ile Gln Gly His His Gly Ala Lys	275	280	285
Gly Glu Arg Gly Glu Lys Gly Glu Pro Gly Val Arg Gly Ala Ile	290	295	300

Sequence Listing

Gly Ser Lys Gly Glu Ser Gly Val Asp Gly Leu Met Gly Pro Ala	305	310	315
Gly Pro Lys Gly Gln Pro Gly Asp Pro Gly Pro Gln Gly Pro Pro	320	325	330
Gly Leu Asp Gly Lys Pro Gly Arg Glu Phe Ser Glu Gln Phe Ile	335	340	345
Arg Gln Val Cys Thr Asp Val Ile Arg Ala Gln Leu Pro Val Leu	350	355	360
Leu Gln Ser Gly Arg Ile Arg Asn Cys Asp His Cys Leu Ser Gln	365	370	375
His Gly Ser Pro Gly Ile Pro Gly Pro Pro Gly Pro Ile Gly Pro	380	385	390
Glu Gly Pro Arg Gly Leu Pro Gly Leu Pro Gly Arg Asp Gly Val	395	400	405
Pro Gly Leu Val Gly Val Pro Gly Arg Pro Gly Val Arg Gly Leu	410	415	420
Lys Gly Leu Pro Gly Arg Asn Gly Glu Lys Gly Ser Gln Gly Phe	425	430	435
Gly Tyr Pro Gly Glu Gln Gly Pro Pro Gly Pro Pro Gly Pro Glu	440	445	450
Gly Pro Pro Gly Ile Ser Lys Glu Gly Pro Pro Gly Asp Pro Gly	455	460	465
Leu Pro Gly Lys Asp Gly Asp His Gly Lys Pro Gly Ile Gln Gly	470	475	480
Gln Pro Gly Pro Pro Gly Ile Cys Asp Pro Ser Leu Cys Phe Ser	485	490	495
Val Ile Ala Arg Arg Asp Pro Phe Arg Lys Gly Pro Asn Tyr	500	505	

<210> 5
 <211> 2865
 <212> DNA
 <213> Homo sapiens

Sequence Listing

<220>

<223> alpha 1 chain collagen

<400> 5

```

atggctcact atattacatt tctctgcatg gttttggtgc tgcttcttca gaattctgtg 60
ttagctgaag atggggaagt aagatcaagt tgctgtactg ctccgacaga tttagttttc 120
atcttagatg gctcttatag tgttggccca gaaaactttg aaatagtga aaagtggctt 180
gtcaatatca caaaaaactt tgacataggg ccgaagttaa ttcaagttgg agtggttcaa 240
tatagtgact accctgtgct ggagattcct ctcggaagct atgattcagg agaacatttg 300
acggcagcag tggaatccat actctactta ggaggaaaca caaagacagg gaaggccatc 360
cagtttgccg tcgattacct ttttgccaag tcctcacgat ttctgactaa gatagcagtg 420
gtacttacgg atggcaaata ccaagatgac gtcaaggatg cagctcaagc agcaagagat 480
agtaagataa cattatattg catttggtgtt gggttcagaaa cagaagatgc cgaacttaga 540
gctattgcca acaagccttc gtctacttat gtgttttatg tggaagacta tattgcaata 600
tccaaaataa gggaagtgat gaagcagaaa ctttgtgaag aatctgtctg tccaacacga 660
attccagtg cagctcgtga tgaaagggga ttgatattc ttttaggttt agatgtaaata 720
aaaaaggtta agaaaagaat acagctttca ccaaaaaaga taaaaggata tgaagtaaca 780
tcaaaagttg atttatcaga actcacaagc aatgttttcc cagaaggctt tcctccatca 840
tatgtatttg tgtctactca aagatttaaa gtcaagaaaa tttgggattt atggagaata 900
ttaactattg atggaaggcc acaaatagca gttaccttaa atgggtgtgga caaaatctta 960
ttatttacia caaccagcgt aattaatggc tcacaagtgg ttacctttgc taaccctcaa 1020
gttaagacgt tgtttgatga aggctggcac caaattcgtc tcttagtaac agaacaagat 1080
gtgactttgt atattgatga ccaacaaatt gaaaacaagc ccttacatcc agtttttaggg 1140
atcttgatca atgggcaaac ccaattgga aaatattctg gaaaagaaga aactgttcag 1200
tttgatgtcc aaaagttgcg aatctactgt gaccagaaac agaacaaccg ggagacagca 1260
tgtgagattc ctggattttg ccttaatggc ccagtgatg taggttcaac tccagctccc 1320

```

Sequence Listing

tgtatttgtc ctccgggaaa accaggactt caaggcccca aaggtgaccc tggactgcct 1380
 gggaaccctg gctaccctgg acaacctggt caagatggta agcctggata tcaggggaatt 1440
 gcagggacac caggtgttcc aggatctcca ggaatacaag gagctcgagg actaccaggt 1500
 tacaaaggag aaccagggcg agatggtgac aagggtgatc gtggacttcc tggttttcct 1560
 gggcttcatg gcatgccagg atcaaagggt gaaatgggtg ccaaaggaga caaaggatca 1620
 cctggatttt atggcaaaaa ggggtgcaaaa ggtgaaaagg ggaatgctgg cttccctggc 1680
 ctccctggac ctgctggaga accaggaaga catggaaagg atggattaat gggtagtccc 1740
 ggtttcaagg gagaagcagg atcccctggt gctccggggc aggatggaac acggggagag 1800
 cctggaatcc caggatttcc tggaaaccga ggattaatgg gccaaaaggg agaaattggg 1860
 cctccaggac agcaaggaaa aaaaggagcc ccagggatgc ctggtttaat gggaagcaat 1920
 ggctcaccag gccagcctgg aacaccggga tctaaggga gcaaaggatga acctggaatt 1980
 caagggatgc ctggggcttc tgggctcaag ggagaaccag gagcaacggg tccccagga 2040
 gaaccaggat acatggggtt acccgggatt caaggaaaaa agggggacaa aggaaatcaa 2100
 ggtgaaaaag gtattcaggg tcaaaaggga gaaaatggaa gacagggaat tccagggcaa 2160
 cagggaattc aaggccatca tgggtcaaaa ggagagagag gtgaaaaggg agaacctggt 2220
 gtccgaggtg ccattggatc aaaaggagaa tctgggggtg atggcttgat ggggcccgc 2280
 ggtcctaagg ggcaacctgg ggatccaggt cctcagggac ccccaggttt ggatgggaag 2340
 cccggaagag agttttcaga acaatttatt cgacaagttt gcacagatgt aataagagcc 2400
 cagctaccag tcttacttca gagtgaaga attagaaatt gtgatcattg cctgtcccaa 2460
 catggctccc cgggtattcc tgggccacct ggtccgatag gccagaggg tcccagagga 2520
 ttacctggtt tgccaggaag agatggtggt cctggattag tgggtgtccc tggacgtcca 2580
 ggtgtcagag gattaaagg cctaccagga agaatgggg aaaaaggag ccaagggttt 2640
 gggtatcctg gagaacaagg tcctcctggt ccccagggtc cagagggccc tcctggaata 2700
 agcaaagaag gtccctcagg agaccaggt ctccctggca aagatggaga ccatggaaaa 2760

Sequence Listing

cctggaatcc aagggcaacc agggcccccga ggcactctgcg acccatcact atgttttagt 2820
gtaattgcca gaagagatcc gttcagaaaa ggaccaaact attag 2865